

# Comparison of Segmentation based on Threshold and K-Means Method

R.Spurgen Ratheash, Dr.M.Mohmed Sathik

Department of Information Technology, Principal  
Sadakathullah Appa College, Tirunelveli  
Tamil Nadu - India

## ABSTRACT

In MRI brain images segmentation, extraction and detection of tumor infected area from the basic brain image properties are the primary, tedious and time taking process. The accuracy of separation the tumor area is based on the experience of clinical experts or radiologists. So, we need computer aided technology to overcome those limitations. In this study, we do automatic methods to reduce the complexity and improve the performance of MRI brain image segmentation. We have investigated many algorithms are available in medical imaging research area such as K-means clustering, Threshold technique, FCMeans, Watershed and Hierarchical Clustering (WHC) and so on. The proposed method compares Threshold technique and K-means clustering technique. The experimental results of proposed techniques have been evaluated and validated for performance and quality analysis on magnetic resonance brain images, based on segmented area, min and major axis and process time for the segmentation. The experimental results achieved more accuracy, less running time and high resolution.

**Keywords:-** Threshold technique, k-means clustering

## I. INTRODUCTION

The uncontrolled growth of cancerous cells in the brain is called as tumours. The brain cells are identified by Benign or Malignant. Malignant is an active cancerous cells with rapid growth in the brain. The Benign cells are not the dangerous cancerous cells. The Benign cells can be converted into Malignant cells but the Malignant cells never become Benign cells.<sup>[1]</sup> This study about the segmentation of abnormal brain cells among normal brain properties such as Gray Matter (GM), White Matter (WM), and CerebroSpinalFluid (CSF) in magnetic resonance (MR) images using Threshold technique<sup>[2]</sup> and K-means Clustering technique.

The Digital image segmentation is employed automatic detection of brain tumor from MRI brain imaging modalities, Segmentation is necessary and important step in image analysis; it is a process of separating an image into different regions, blocks or clusters sharing common and identical properties, such as contrast, patterns of pixels, and distance around the boundary of the region, and gray level.

## II. LITERATURE SURVEY

Automatic identifying and extraction of brain tumor has proposed by the techniques like Threshold, K-Means Clustering, Fuzzy Clustering Means (FCM), Pulse Couple Neural Network(PCNN) algorithm, Expectation Maximization (EM) segmentation algorithm, Watershed and Hierarchical Clustering (WHC) algorithm, support vector machine (SVM), artificial neural network (ANN) algorithm<sup>[3]</sup>. The above literature survey has revealed that some of the techniques are invented to obtain segmentation the brain area from the skull area; some of

the techniques are invented to obtain feature extraction and some of the techniques are invented to obtain classification only<sup>[4]</sup>. Threshold Technique based on image intensity and K-means clustering algorithm based on the clusters<sup>[5]</sup> in MRI brain images, the K-means clustering method gives an effective segmentation of tumor region. This analysis on combined approach could not be conducted in any published literature.

## III. THRESHOLD

In Threshold technique is based on histogram to identify the infected areas by deep and sharp valley between two peaks representing objects and background respectively.

The threshold can be chosen at the bottom of this valley. However, for most MR images, it is often difficult to detect the valley bottom precisely when the valley is flat and broad, imbued with noise, or when the two peaks are extremely unequal in height, often producing no traceable valley<sup>[6,7]</sup>. The threshold method can choose the value and separate the object from its background.

Let the pixels of a given picture be represented in  $L$  gray levels  $[1, 2, \dots, L]$ . The number of pixels at level  $i$  is denoted by  $n_i$  and the total number of pixels by  $N = n_1 + n_2 + \dots + n_L$ . In order to simplify the discussion, the gray-level histogram is normalized and regarded as a probability distribution:

$$P_i = n_i / N, P_i \geq 0, \sum_{i=1}^L P_i = 1 \quad (1)$$

Now suppose that we dichotomize the pixels into two classes  $C_0$  and  $C_1$  (background and objects, or vice

versa) by a threshold at level  $K$ ;  $C_0$  denotes pixels with levels  $[1, \dots, k]$ , and  $C_1$  denotes pixels with levels  $[k+1, \dots, L]$ . Then the probabilities of class occurrence and the class mean levels, respectively, are given by

$$\omega_0 = P_r(C_0) \tag{2}$$

$$\omega_1 = P_r(C_1) \tag{3}$$

$$\mu_0 = \sum_{i=1}^k i P_r(i / C_0) \tag{4}$$

$$\mu_1 = \sum_{i=k+1}^L i P_r(i / C_1) \tag{5}$$

where

$$\alpha(k) = \sum_{i=1}^k P_i \tag{6}$$

and

$$\omega(k) = \sum_{i=1}^k i P_i \tag{7}$$

are the zeroth and the first-order cumulative moments of the histogram up to the  $k$ th level, respectively, and

$$\mu_r = \mu(L) = \sum_{i=1}^L i P_i \tag{8}$$

is the total mean level of the original picture. We can easily verify the following relation for any choice of  $k$

$$\omega_0 \mu_0 + \omega_1 \mu_1 + \dots + \mu_r, \quad \omega_0 + \omega_1 = 1 \tag{9}$$

The Class variances are given by

$$\sigma_0^2 = \sum_{i=1}^k (i - \mu_0)^2 P_r(i / C_0) \tag{10}$$

$$\sigma_1^2 = \sum_{i=k+1}^L (i - \mu_1)^2 P_r(i / C_1) \tag{11}$$

These require second-order cumulative moments (statistics).

In order to evaluate the “goodness” of the threshold (at level  $k$ ), we shall introduce the following discriminant criterion measures (or measures of class separability) used in the discriminant analysis [5]:

$$\lambda = \sigma_B^2 / \sigma_W^2, \quad K = \sigma_I^2 / \sigma_W^2, \quad \eta = \sigma_B^2 / \sigma_I^2 \tag{12}$$

where

$$\sigma_W^2 = \omega_0 \sigma_0^2 + \omega_1 \sigma_1^2 \tag{13}$$

$$\begin{aligned} \sigma_B^2 &= \omega_0 (\mu_0 - \mu_r)^2 + \omega_1 (\mu_1 - \mu_r)^2 \\ &= \omega_0 \omega_1 (\mu_1 - \mu_0)^2 \end{aligned} \tag{14}$$

(due to (9)) and

$$\sigma_I^2 = \sum_{i=1}^L (i - \mu_r)^2 P_i \tag{15}$$

are the within-class variance, the between-class variance, and the total variance of levels, respectively. Then our problem is reduced to an optimization problem to search for a threshold  $k$  that maximizes one of the object functions (the criterion measures) in (12).

This standpoint is motivated by a conjecture that well-threshold giving the best separation of classes in gray levels would be the best threshold.

The discriminant criteria maximizing  $\lambda$ ,  $k$  and  $\eta$  respectively, for  $k$  are, however, equivalent to one another; e.g.,  $k = \lambda + 1$  and  $\eta = \lambda / (\lambda + 1)$  in terms of  $\lambda$ , because the following basic relation always holds:

$$\sigma_W^2 + \sigma_B^2 = \sigma_I^2 \tag{16}$$

It is noticed that  $\sigma_W^2$  and  $\sigma_B^2$  are functions of threshold level  $k$ , but  $\sigma_I^2$  is independent of  $k$ .<sup>[8]</sup> It is also noted that  $\sigma_W^2$  is based on the second-order statistics (class variances), while  $\sigma_B^2$  is based on the first-order statistics (class means)<sup>[9]</sup>. Therefore,  $\eta$  is the simplest measure with respect to  $k$ . Thus we adopt  $\eta$  as the criterion measure to evaluate the “goodness” (or separability) of the threshold at level  $k$ .

The optimal threshold  $k^*$  that maximizes  $\eta$ , or equivalently  $\sigma_B^2$  maximizes is selected in the following sequential search by using the simple cumulative quantities (6) and (7), or explicitly using (2)-(5):

$$\eta(k) = \sigma_B^2(k) / \sigma_I^2 \tag{17}$$

$$\sigma_B^2(k) = \frac{[\mu_r \omega(k) - \mu(k)]^2}{\omega(k)[1 - \mu(k)]} \tag{18}$$

and the optimal threshold  $k^*$  is

$$\sigma_B^2(k^*) = \max_{1 \leq k < L} \sigma_B^2(k) \tag{19}$$

From the problem, the range of  $k$  over which the maximum is sought can be restricted to

$$\begin{aligned} S^* &= \{k; \omega_0 \omega_1 = \alpha(k)[1 - \alpha(k)] > 0 \\ &\quad \text{or} \\ &\quad 0 < \alpha(k) < 1\} \end{aligned}$$

We shall call it the effective range of the gray-level histogram. From the definition in (14), the criterion measure  $\sigma_B^2$  (or  $\eta$ ) takes a minimum value of zero for such  $k$  as  $k \in S - S^* = \{k; \alpha(k) = 0 \text{ or } 1\}$  (i.e., making all pixels either  $C_1$  or  $C_0$ , which is, of course, not our concern) and takes a positive and bounded value for  $k \in S^*$ . It is, therefore, obvious that the maximum always exists.

Threshold algorithm applied on the dataset which is in the DICOM images that has converted into JPEG images then apply the Threshold level of 75. The value of threshold will set depends on the contrast of an image. The figure1 shows segment the tumor alone from the MRI brain tumor image.

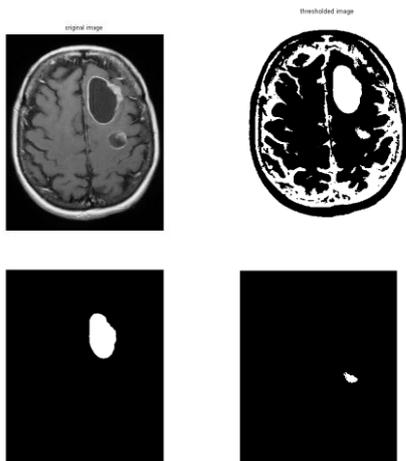


Figure 1 Segmented tumor area by Threshold technique

#### IV. THE K-MEANS ALGORITHM

Since the simplest form of K-means clustering algorithm, it is used by most of the researchers in the field of data mining. The process of k-Means follows eminent and simple way to classify a given data set by means of a certain number of clusters. The K-means clustering is a method used to divide an image as  $n$  patterns  $\{x_1, x_2, \dots, x_n\}$  in  $d$  dimensional space into  $k$  clusters (assume  $k$  clusters).<sup>[10]</sup> The result is a set of  $k$  clusters based on  $k$  centres, each of which is located at the centroid of the separated dataset. This algorithm can be shortened in the following steps:

The steps involved in clustering the MRI brain images by k-Means algorithm are given below.

- Step 1: Insert the original images as input.
- Step 2: Convert the fetched MRI DICOM format file into .JPG
- Step 3: Cluster dataset images.
- Step 4: Find out the 'k' in image by algorithm itself.
- Step 5: Get the clustered objects.

The process on clustered images are given below.<sup>[11]</sup>

- Step 1: Give the number of cluster value as  $k$ .
  - Step 2: Randomly choose the  $k$  cluster centres.
  - Step 3: Calculate mean or centre of the cluster.
  - Step 4: Calculate the distance between each pixel to each cluster centre.
  - Step 5: If the distance is near to the centre then move to that cluster.
  - Step 6: Otherwise move to next cluster.
  - Step 7: Re-estimate the centre.
  - Step 8: Repeat the process until the centre doesn't move.
- The k-means clustering algorithm fixes the  $k$  value as 5 in MR image.

In this research work, tumor detection by identifying the pixel values in MRI brain images are taken for analysis. The source code is written in MATLAB software. The k-Means algorithm is applied to find the clusters of MRI images by dividing the image into 5 groups. The various stages of images are given in the Figure 2.

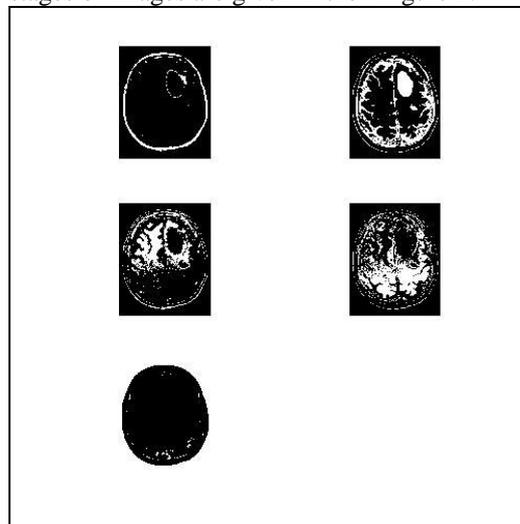


Figure 2. K means algorithm output when  $k=5$

Result of segmented image by K-means clustering in Figure 3.

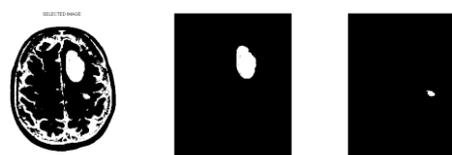
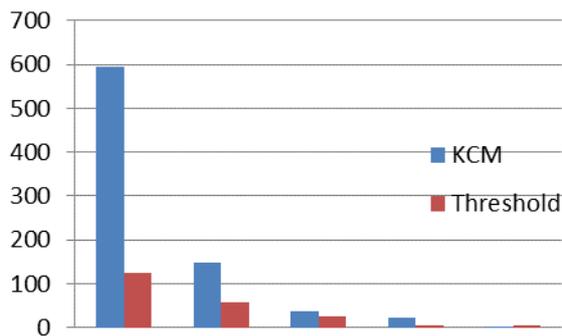


Figure 3 Segmented tumor area by K-means clustering

The figures from figure-1 and figure-3 were calculated and isolated brain tumor from its properties by two methods. The first stage is to determine the area of segmented image. The second stage is to determine the process time, perimeter, minimum axis and maximum axis of the segmented image. Segmented area is automatically calculated from MRI brain images. The segmented images are compare by the following properties.

Methods	Area	Perimeter	Major Axis	Minor Axis	Process Time
KCM	595	147	37	21	3.3252
Threshold	124	57	24	6	5.5813



## V. CONCLUSION

In this paper, a fully automated tumor detection method based on Threshold and K-means clustering techniques are proposed. Threshold technique segments the tumor area by its intensity value in MRI brain images. The K-means clustering gives tumor area by cluster the object which is the methods. The K-means clustering has minimum processing time and also gives an accurate value for the infected area. In future, the entire tumor area identification approach is extendable to 3D to convert into volumetric data. The K-means clustering method is a suitable method to segment brain MR images. This method can also be applied to other medical images e.g., heart or liver MRI.

## REFERENCES

- [1] M. Sezgin and B. Sankur, "Survey over Image Thresholding Techniques and Quantitative Performance Evaluation", *Journal of Electronic Imaging*, Vol. 13, pp. 146-165, 2004.
- [2] N. Otsu, "A Threshold Selection Method from Gray-Level Histogram", *IEEE Trans. Systems Man, and Cybernetics*, Vol. 9, pp. 62-66, 1979.
- [3] Matthew C. Clark "Segmenting MRI Volumes of the Brain With Knowledge- Based Clustering" MS Thesis, Department of Computer Science and Engineering, University of South Florida, 1994
- [4] Dzung L. Pham, Chenyang Xu, Jerry L. Prince, "A Survey of Current Methods in Medical Medical Image Segmentation" Technical Report JHU / ECE 99-01, Department of Electrical and Computer Engineering. The Johns Hopkins University, Baltimore MD 21218, 1998.
- [5] Osama Abu Abbas "Comparison of various clustering algorithms" *The International Arab Journal of Information Technology*, Vol. 5, No. 3, July 2008.
- [6]. S.Xavierarockiaraj, K.Nithya and R.Maruni Devi, "Brain tumour Detection Using Modified Histogram thresholding-Quadrant Approach", *Journal of Computer Applications(JCA)*, Vol. 5, No.1, 2012, pp.21-25.
- [7]. Anam Mustaqeem, ali Javed and Tehseen Fatima, "An Efficient Brai tumour Detection Algorithm using Watershed and Thresholding based segmentation",

- International Journal of Image, Graphics and signal Processing (IJIGSP)*, Vol. 4, No.10, 2012, pp.34-39.
- [8]. Kai Xie, Jie Yang, Z.G.Zhang and Y.M.Zhu, "Semi-automated brain tumour and edema segmentation using MRI.", *European Journal of Radiology*, Vol.56, Issue 1,October 2005, pp.12-19
- [9]. Nikola K. Kasabov, *Foundations of Neural Networks, Fuzzy systems and Knowledge Engineering*, Massachusetts Institute of Technology, 1998, pp.167-473
- [10] Jaehui Park, Sang-goo Lee "Probabilistic Ranking for Relational Databases based on Correlations" *ACM* 2010.
- [11] Kanade BP, Gumaste PP. Brain tumordetection using MRI images. *International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering*. 2015; 3(2):146–50.